NOVEMBER/DECEMBER 2024

23UDA31 — FUNDAMENTALS OF DATA SCIENCE

Time : Three hours Maximum : 75 marks

SHANMUGA INDUSTRIES ARTS & SCIENCE COLLEGE LIBRARY

SECTION A — (10 × 2 = 20 marks)

Answer ALL questions.

1. Define Data Science.
2. State Array Operations.
3. Why is Pandas used?
4. Delineation of Pandas Libraries.
5. What is Text File?
6. Write an Syntax for Reading a File using Readlines ( ).
7. Characterize Data Wrangling.
8. What is Hierarchical Indexing?
9. List Down the types of Plots in Matplotlib.
10. Specify Subplot in note.

SECTION B — (5 × 5 = 25 marks)

Answer ALL questions.

11. (a) Expound main elements of Datafication in Data Science.

Or

(b) Demonstrate Data Science Process Life Cycle.

12. (a) Write a short note on Pandas Data Frame.

Or

(b) Elucidate ReIndexing.

13. (a) Difference between Mongo DB and RDBMS.

Or

(b) Explain JSON data types.

14. (a) Discuss Merge two Pandas Data frames on Index.

Or

(b) Elaborate Concatenating Along an Axis.

15. (a) Compare between Figure and Subplot.

Or

(b) Illustrate Matplotlib Pyplot API.

SECTION C — (3 × 10 = 30 marks)

Answer any THREE questions.

16. Delegate about Numpy Basics in detail.

17. Construct Handling and Filling Missing data in point.

18. Describe Mongo DB working and Features in detail.

19. Discuss about Combining and Merging data sets in brief.

20. Elaborate Histograms and Density Plots.

———————